

I. AMENDMENTS TO THE SPECIFICATION

Kindly replace the paragraph beginning on page 12, line 27, with the following new paragraph:

The second input 114 is for receiving a set of information extraction rules defining the semantic and/or syntactic characteristics ~~of that~~ that the natural language information extraction system 160 is to seek in the sentence. In a variant, the set of information extraction rules may be fixed in the natural language information extraction system 160 such that the system 160 is configured to perform a specific information extraction function. In this variant, the second input 114 may be omitted.

Kindly replace the paragraph beginning on page 14, line 2, with the following new paragraph:

The morphological analyzer 100 has an input for receiving the textual representation of a sentence. The morphological analyzer processes the sentence to determine the possible parts of speech and lemmas, herein referred to as morphological tags, for the words in the sentence and then assigns a most likely morphological tag for each individual word by applying contextual rules based on different classes of ambiguity. The morphological analyzer 100 then locates boundaries of noun phrases. The morphological analyser also has an output for releasing a labeled sequence of word/tag pairs and noun phrases corresponding to the textual representation of the sentence received at the input.

Kindly replace the paragraph beginning on page 26, line 7, with the following new paragraph:

The syntactic processor 102 will be best understood in connection with the flow chart of figure 5. For the purpose of illustration, the parse trees in the parse tree group includes a root node associated to a word (or noun phrase) of the sentence. For each sentence, a parse tree group comprising a set of parse trees is obtained by first creating an empty parse tree group in step 500. Following this at step 502, for each word in the sentence, a parse tree with one (1) element is created and inserted into the parse tree group at step 504. At step 506, binary dependency rules are applied to each parse tree newly inserted in the parse tree group. The binary dependency rules applied attempt to create a new parse tree by combining the newly inserted parse tree with any other parse tree having a root node associated to a word that is a precursor or a successor of the word associated to the root node of the newly inserted tree. The new parse tree includes the ~~words~~ words of the two parse trees as well as a data element indicating the dependency relationship of the two parse trees.

Kindly replace the paragraph beginning on page 31, line 29, with the following new paragraph:

Yet another advantage of the syntactic processor in accordance with the spirit of the invention is that it permits maintaining information about the relationships of the words in the sentence even if the entire sentence could not be resolved beyond a certain point (because of a grammatical incongruity for example). In other words, even ~~if a the~~ if the parsing of a sentence could not be completed beyond a certain point, the parse trees for the portions of the sentence are preserved, thus allowing information to be obtained in spite of this failure to complete the parse.

Kindly replace the paragraph beginning on page 32, line 14, with the following new paragraph:

In a specific example of implementation, in addition to the parse tree group received from the syntactic processor 102, the information extraction unit 108 further receives from the semantic analyzer 104 noun phrases associated to respective semantic types and from the second input 114 information extraction rules. The information extraction rules ~~defines~~ define the semantic and/or syntactic characteristics ~~of that~~ that the information extraction unit 108 is to seek in the ~~parse~~ parse tree group. The information extraction rules may be stored in the information extraction unit 108 such that the latter is configured to perform a specific information extraction function. In this variant, the second input 114 may be omitted.

Kindly replace the paragraph beginning on page 33, line 6, with the following new paragraph:

In a specific example of implementation, the set of information extraction rules are in the form of semantic frames containing information related to:

- a) the name ~~or the~~ of the resulting information record that will be generated when a match to a given parse tree pattern occurs;
- b) the specific word OR noun phrase semantic type to be found as a head of a parse tree, and
- c) a list of dependants to be found for this head in the parse tree. Each dependant is either obligatory or optional and is expressed as:
 - i) a dependency name to be found in the parse tree;
 - ii) the name of the field within the information record for this element;
 - iii) an optional noun phrase semantic type for this dependant.